AD_____

Award Number: DAMD17-02-1-0489

TITLE: Computer-Aided Interval Change Analysis of Microcalcifications on Mammograms for Breast Cancer Detection

PRINCIPAL INVESTIGATOR: Lubomir Hadjiiski, Ph.D.

CONTRACTING ORGANIZATION: University of Michigan
Ann Arbor, Michigan 48109-0904

REPORT DATE: July 2006

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
01-07-2006	Annual	1 JUL 2005 - 30 JUN 2006

4. TITLE AND SUBTITLE
Computer-Aided Interval Change Analysis of Microcalcifications on Mammograms for Breast Cancer Detection

5a. CONTRACT NUMBER

5b. GRANT NUMBER
DAMD17-02-1-0489

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Lubomir Hadjiiski, Ph.D.

E-Mail: lhadjisk@umich.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Michigan
Ann Arbor, Michigan 48109-0904

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT: The goal of this project is to develop a computer-aided diagnosis (CAD) system for automatic interval change analysis of microcalcification clusters on mammograms. Based on our regional registration method and a search program cluster candidates were detected within the local area on the prior. The cluster on the current image is then paired with the candi-dates to form true (TP-TP) or false (TP-FP) pairs and a correspondence classifier is designed to reduce the (TP-FP). A temporal classifier (TC) based on current and prior information is used if a cluster is detected in the prior, and a current classifier (CurC) based on current information alone is used if no prior cluster is detected. For the TC an LDA, SVM and NN were used. 175 temporal pairs of mammograms were used for evaluation. The registration stage identified 85% (149/175) of the TP-TP pairs with 15 false matches within the 164 image pairs that had detected clusters. The TC based on LDA, SVM and NN achieved a test Az of 0.83, 0.82, 0.84, respectively, for the 164 pairs for classifying the clusters as malignant or benign. For the 11 clusters without detection on the prior, the test Az by the CurC was 0.72. Four radiologists participated in pilot observer study using our CAD. The average Az in estimating the likelihood of malignancy was 0.70 without CAD and improved to 0.77 with CAD(p=0.04).

15. SUBJECT TERMS
Breast Cancer, Computer-aided diagnosis, Screening, Classification, Image Analysis

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	33	19b. TELEPHONE NUMBER *(include area code)*

(3) Table of Contents

(4) Introduction

Treatment of breast cancer at an early stage can significantly improve the survival rate of patients. Mammography is currently the most sensitive method for detecting early breast cancer, and it is also the most practical for screening. Although general rules for differentiation between malignant and benign lesions exist, in clinical practice, approximately only 15-30% of cases referred to surgical biopsy are actually malignant. A number of research groups are in the process of developing computer-aided diagnosis (CAD) methods which can provide a consistent and reproducible second opinion to the radiologist for the detection and classification of breast abnormalities.

Radiologists routinely compare mammograms from a current examination with those obtained in previous years, if available, for identifying interval changes, detecting potential abnormalities, and in evaluating breast lesions. It is widely accepted that interval changes in mammographic features are very useful for both detection and classification of abnormalities. However, CAD techniques that use multiple exams for detection or characterization have not been commonly explored, probably because of the difficulty in the registration of the compressed breast images from different exams. We have been investigating methods for analysis of temporal changes of masses on mammograms to improve detection and classification. To our knowledge, there is no existing CAD technique for registration of microcalcification clusters or classification of microcalcifications based on temporal change information.

The extraction of any meaningful information from a prior mammogram first requires a common frame of reference between the current and prior mammograms. Several complicating factors, such as breast compression difference between current and prior mammograms, energy difference between the two imaging conditions, differences in screen film properties and film processing conditions, and potential changes in breast structures between the two images with patient age, make it difficult to obtain such a frame of reference. On breast images, there are no invariant landmarks (except for the nipple) that can serve as control points in conventional image registration methods to register the two mammograms. In this project, we propose to develop an innovative regional registration method that does not depend on specific control points. We will first approximately align the current and prior mammogram based on maximization of mutual information. Next, we will design a novel approach in which the computer emulates the radiologists' search method in finding corresponding lesions on mammograms. Automated search of microcalcification cluster within the search region on the prior mammogram will be performed. Our current automated microcalcification detection algorithm will provide a basis for this search. However, since the detection is limited to the small search region, the detection can be performed in high resolution and the algorithm parameters can be adjusted to improve the detection sensitivity of the very subtle clusters on the prior mammograms without excessive trade-off in increasing false-positives. A correspondence classifier will be developed to identify the matched pair of clusters on the two mammograms. The image features of the corresponding microcalcification clusters can then be automatically extracted and feature measures characterizing interval changes derived. A classification scheme to differentiate malignant and benign clusters using the interval change information will be developed. This computerized interval change analysis will be an important component of a CAD system for mammographic interpretation.

This project aims at developing a novel interval change analysis scheme to improve the accuracy of CAD. We will investigate the problem of classifying microcalcifications as malignant or benign based on temporal changes in mammographic features using a combination

of computer vision, automated feature extraction, statistical classification, and artificial intelligence techniques. We hypothesize that the use of temporal information would improve the ability of CAD to offer an accurate and objective second opinion to radiologists which, in turn, would increase the positive predictive value of mammography, reduce the number of benign biopsies, and hence reduce both cost and patient morbidity. If integrated in a complete CAD system, the algorithms to be developed in this project may also increase the efficacy of mammography for early detection of breast cancer.

(5) Body

In the fourth year (7/1/05-6/30/06) of this grant, we have performed the following studies:

(A) Database collection of malignant and benign breast microcalcification cases that have multiple examinations (Task 1)

We continued collecting the data set for this study from the files of patients who had undergone biopsy at the University of Michigan. The mammograms are scanned and the images are saved in our storage device using automated graphic user interface developed in our laboratory. Additionally the film information is recorded in a Microsoft Access database. Temporal pairs of images were obtained. The current mammogram of each temporal pair exhibited a biopsy-proven mass. We scan both cranio-caudal (CC) and mediolateral-oblique (MLO) views. The mammograms were digitized with a LUMISCAN 85 laser scanner at a pixel resolution of 0.05 mm x 0.05 mm and with 12-bit resolution.

While the regional registration technique can be used for determining a corresponding structure or region for any structure (both normal tissues and masses) in the breast, in this study we are analyzing its accuracy on biopsy-proven masses alone. The location of the mass on the current mammogram is identified by an Mammography Quality Standards Act (MQSA)-approved radiologist experienced in breast imaging using an interactive image analysis tool on a UNIX workstation. To provide the ground truth for evaluation of the computerized method, the radiologist manually identifies the corresponding region on the prior mammogram. Bounding polygons enclosing the microcalcification cluster on the current mammogram and the corresponding object on the prior mammogram are provided by the radiologist for each case. Each microcalcification cluster as well as the corresponding structure on the prior mammogram are rated for its visibility on a scale of 1 to 10, where the rating of 1 corresponded to the most visible category. The size of the microcalcification cluster on the current mammogram as well as the size of the corresponding structure on the prior mammogram are also measured by the radiologist. The parenchymal density is rated based on the Breast Imaging Reporting and Data System (BI-RADS) lexicon.

(B) Development of feature measures and temporal classifier for characterization of temporal changes in microcalcification clusters. (Task 6)

In the past year of the project we made a major progress in the direction of designing an automatic CAD system for characterization of temporal cases on malignant and benign. This is very novel and unique CAD system which includes an automatic registration of the corresponding

current and prior clusters and then classifying them by a temporal classifier on malignant and benign (Figure 1). For this purpose the feature extraction and classification were carried out with the clusters and individual microcalcifications obtained automatically from the registration stage [1, 6, 9]. The design of an automatic CAD also presents challenges. The automatically detected cluster locations on the temporal pairs may deviate from the optimal locations as selected by expert radiologists. This will introduce "noise" to the extracted features and make the classification task more difficult. In addition for some cases there is no detection on prior mammogram and only current information can be used.

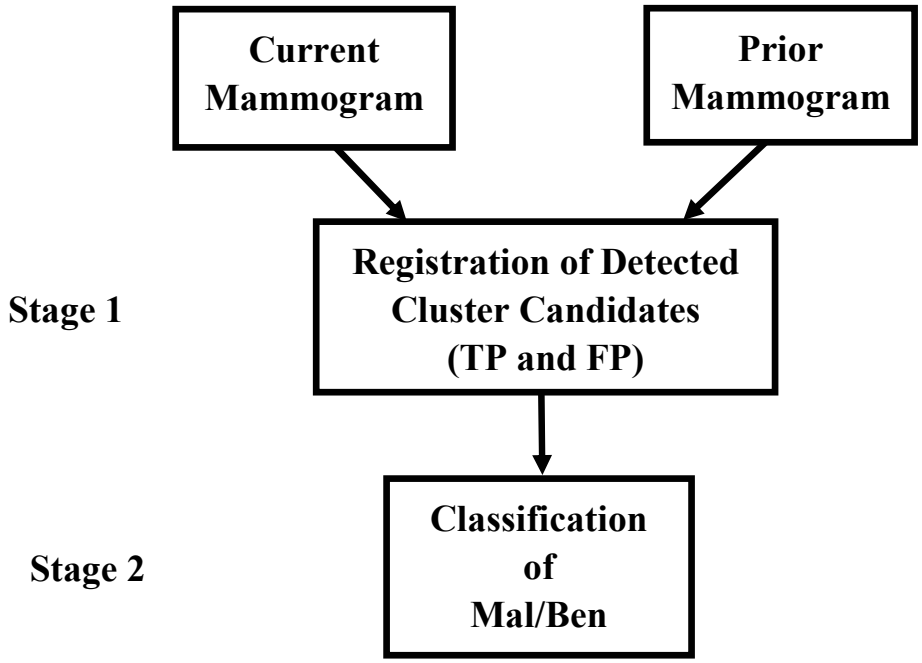

Figure 1. Block-diagram of the registration – classification CAD for temporal microcalcification clusters.

Feature extraction and definition of difference features

In this study, a new classification scheme using interval change information was developed to classify mammographic microcalcification clusters as malignant and benign [9] (Figure 2).

From each automatically detected cluster, 20 run length statistic texture features (RLSF) and 21 morphological (Mo) features were extracted [7]. Additionally, 78 SGLD [2] and 64 GLDS

[3, 4] texture features were also extracted. All texture features (RLSF, SGLD, and GLDS) were extracted from automatically detected cluster locations. The morphological features were extracted from the automatically detected microcalcifications within the automatically detected cluster locations.

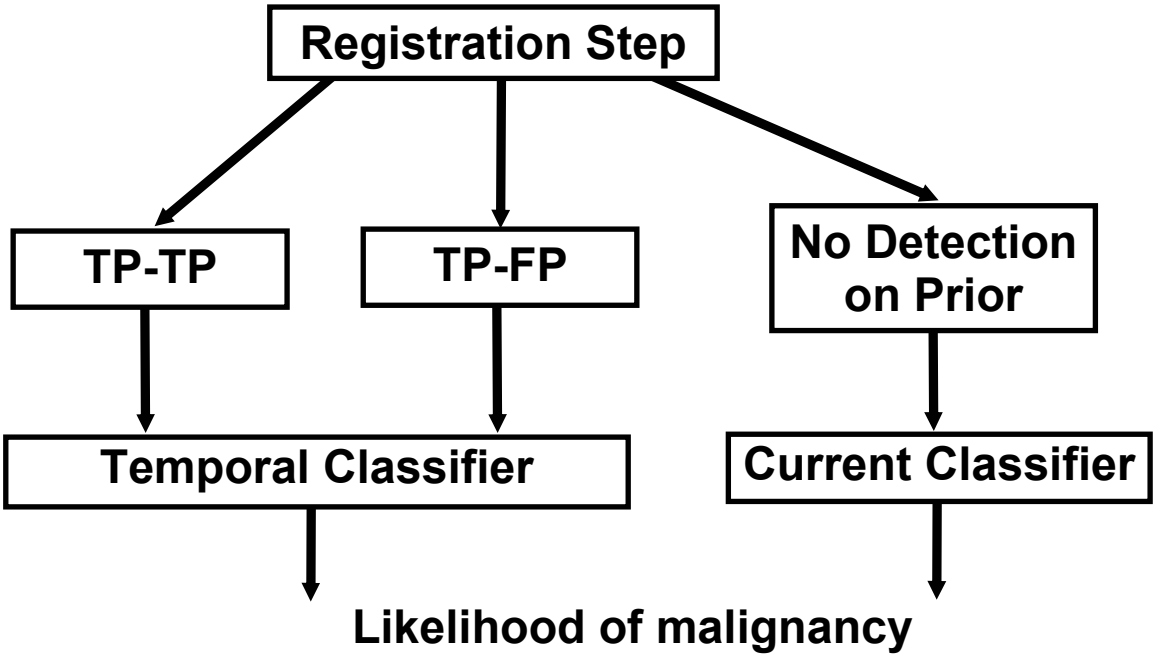

Figure 2. Classification of temporal pairs of microcalcification clusters on malignant and benign combining temporal classifier for the TP-TP and TP-FP pairs and current classifier for the cases without detection on prior.

Twenty difference RLSF were obtained by subtracting a prior RLSF from the corresponding current RLSF. We have designed a new feature, the ratio feature, which is defined as the ratio between current and prior feature [5, 9]. We have obtained 21 Mo ratio features. In addition we used current GLDS features.

Table 1. The feature vector used for the temporal classifier.

Generated feature type	Ratio features	Difference features	Current features
Features	Mo	RLSF	GLDS
Number of features	21	20	64

The feature space consisted of the Mo ratio features, the difference RLSF, as well as the current GLDS features (Table 1).

We are continuing to investigate the possibilities to define and extract additional sets of features in order to improve the classification accuracy.

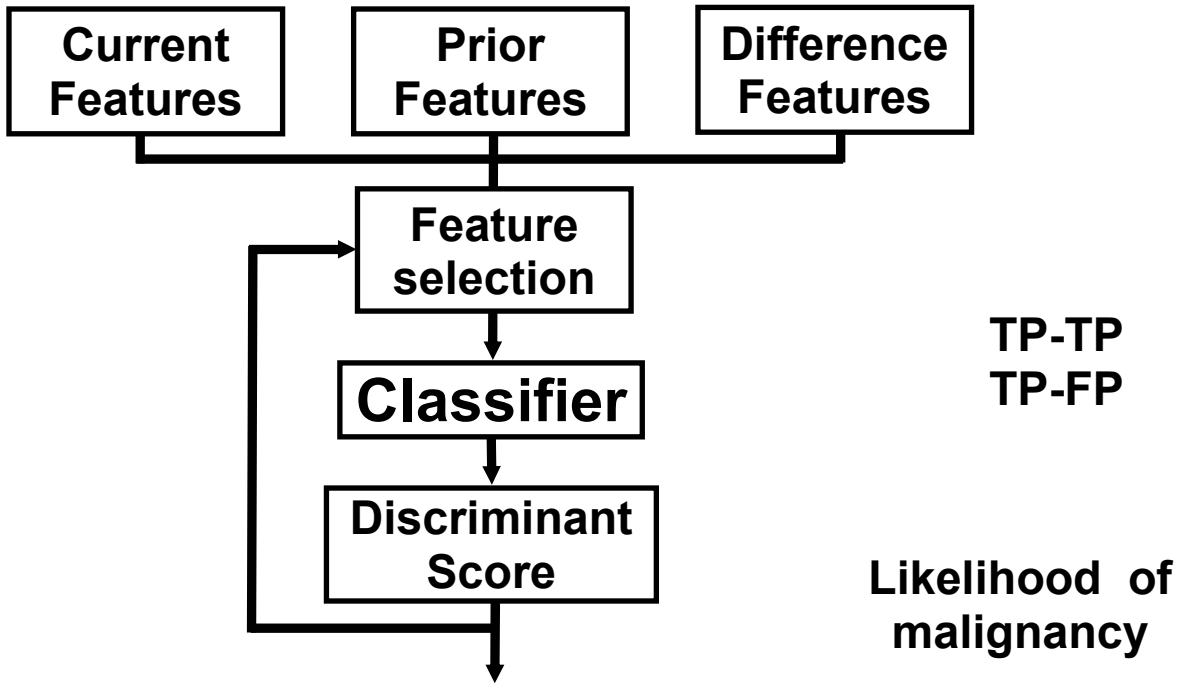

Figure 3. Temporal classifier for classification of temporal pairs of microcalcification clusters on malignant and benign.

Classifier design

When we are designing the classification system for classification of microcalcification clusters on malignant and benign we have to consider the fact that the microcalcification detection algorithm may not be able to detect all clusters on prior. If the cluster is very subtle the automatic microcalcification detection for the identification of mirocalcification clusters may not be able to detect it. In this case two types of classifiers are designed (Figure 2). A temporal classifier (Figure 3) based on current and prior information is used if a cluster is detected in the prior, and a current classifier (Figure 6) based on current information alone is used if no prior cluster is detected. The temporal classifier is trained to classify the true (TP-TP) and false (TP-FP) pairs. The automatically detected cluster locations on the temporal pairs may deviate from the optimal locations as selected by expert radiologists. This will introduce "noise" to the extracted features and make the classification task more difficult.

In the past year of the project we have concentrated our efforts on using different types of classification methods in the temporal classifier in order to classify the temporal pair on malignant and benign. We have used a linear discriminant analysis classifier (LDA), support vector machine classifier (SVM) [8] and a multilayer perceptron neural network with backpropagation training (NN). A stepwise feature selection with floating window of 3 were used to select the most useful feature subsets and to merge the feature into a discriminant score (Figure 3). A leave-one-case-out resampling scheme was used for feature selection and to train and to test the classifiers.

In addition a current classifier was trained using the current images from the temporal pairs (the cases that has no detection on prior) and tested on the cases that have not detection on prior (Figure 6). LDA classifier was used in this task. The classification accuracy was analyzed by receiver operating characteristic (ROC) methodology.

In this study, 175 serial pairs containing biopsy-proven calcification clusters were used. On the priors, the radiologist rated 12 of the 175 clusters as not visible and the subtlety of 18 clusters as 9 or 10 on a scale of 10.

At the first stage of the system (Figure 1), 85% (149/175) of the TP-TP pairs were identified with 15 false matches within the 164 image pairs that had computer detected clusters on the priors. Therefore 164 temporal image pairs (149 TP-TP and 15 TP-FP), of which 49 were malignant, were used for classification as malignant and benign. At the second stage (Figure1), the SVM, LDA and NN temporal classifiers were used for the classification of the 164 cluster temporal pairs.

For each of LDA and SVM (both radial and dot kernel) temporal classifiers a feature selection was performed. Feature selection for NN is a very competitively intensive process and because of this for the NN temporal classifier we used the best feature set obtained for the LDA temporal classifier.

LDA temporal classifier

When an LDA temporal classifier was used the best result was obtained for an average of 7 features selected (Table 2). The selected features included 1 difference RLS feature, 4 morphological ratio features and 2 GLDS features from the current image. All the features were consistently selected for all the training partitions. The LDA temporal classifier achieved test A_z of 0.83±0.03 (Table 4). The A_z results for the LDA temporal classifier with different number of selected features is presented in Figure 4.

Table 2. Selected features for the combined classifier for classification on malignant and benign of automatically detected clusters in current and prior mammograms.

Classifier	Feature type	Ratio features	Difference features	Current features
Temporal classifier LDA, NN	Mo	4		
	RLSF		1	
	GLDS			2
Temporal classifier SVM (dot kernel)	Mo	3		
	RLSF		1	
	GLDS			3
Temporal classifier SVM (radial kernel)	Mo	3		
	RLSF		1	
	GLDS			3

SVM temporal classifier

SVMs with dot kernel, radial kernel and tanh kernel (neural kernel) [8] were studied [10]. The training and testing results are presented in Table 3.

The SVM temporal classifier with dot kernel achieved training A_z of 0.87±0.03 and test A_z of 0.82±0.03. The selected features for the dot kernel included 1 difference RLS feature, 3 morphological ratio features and 3 GLDS features from the current image. The performance of the SVM with dot kernel for the different number of selected features is presented in Figure 4. For the small number of features the performance of the SVM with dot kernel was close to LDA. However for the larger number of features the LDA slightly over-performed the SVM with dot kernel.

The SVM temporal classifier with radial kernel achieved training A_z of 0.86±0.03 and test A_z of 0.81±0.03 (Table 3). For this kernel the selected features were also 1 difference RLS feature, 3 morphological ratio features and 3 GLDS features from the current image (Table 2). The A_z values of the SVM temporal classifier with radial kernel were consistently smaller than that for the SVM with dot kernel for the different number of features. (Figure 4).

The SVM temporal classifier with tanh kernel (neural kernel) achieved training A_z of 0.86±0.03 and test A_z of 0.82±0.03 (Table 3). However, it required 83 support vectors and it was slower to train. Because of that we did not perform feature selection used the best feature set we had from the LDA classifier.

The SVM with dot kernel performed the best compared to the SVMs with other kernels. It also required the smallest number of support vectors.

NN temporal classifier

A NN with 7 input nodes, different number of neurons in the hidden layer, and 1 output node was used. The NN achieved the best result (A_z = 0.84±0.03) for 8 hidden nodes (Figure 5). The performance of the NN temporal classifier as a function of the number of the hidden neurons is presented in Figure 5. For most of the NN structures the test A_z was 0.83 or lower. For larger number of hidden neurons we have observed slight overtraining. For 8 hidden nodes there was a pick in A_z, which corresponded to the best result.

Table 4 summarizes the results of using different classification methods for the temporal classifier. The difference in A_z between any one of the classifiers did not achieve statistical significance. LDA classifier revealed good stable performance with very simple structure of the classifier. This was the reason to be selected for the observer study experiment.

In comparison, the MQSA radiologist achieved an A_z of 0.72±0.04 for the 164 temporal pairs.

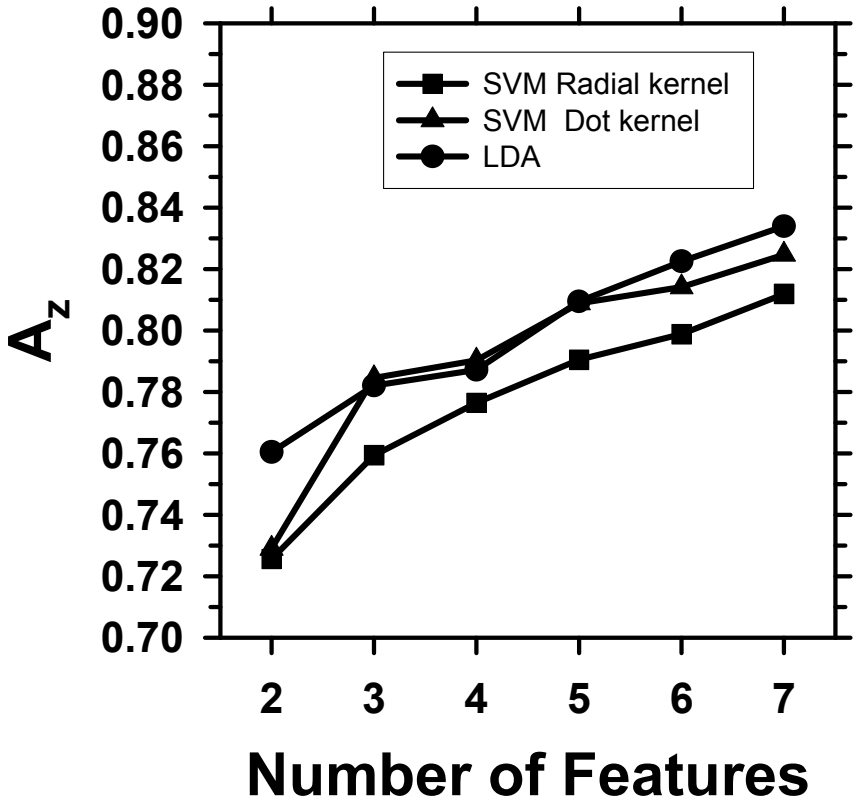

Figure 4. Test Az results for SVM (radial and dot kernels) and LDA with different number of selected features.

The difference RLS texture features, morphological ratio features and GLDS current features were useful for identification of malignancy in temporal pairs of mammograms. The information on the prior image improved characterization of the microcalcification clusters: 5 out of the 7 selected features contained prior information for the LDA and NN; 4 out of the 7 selected features contained prior information for the SVM.

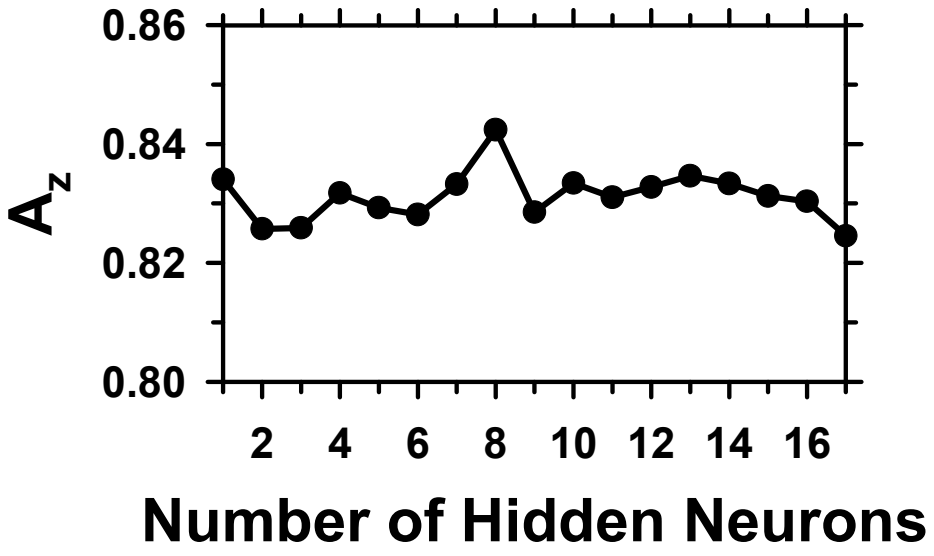

Figure 5. Test results for NN with different number of neurons in the hidden layer.

Table 3. SVM classification results based on 7 selected features.

SVM Type	Training A_z	Test A_z	Number SV
Dot Kernel	0.87±0.03	0.82±0.03	78
Radial Kernel	0.86±0.03	0.81±0.03	89
Tanh Kernel	0.86±0.03	0.82±0.03	83

Table 4. Overall results for the LDA, SVM, NN classifers.

	Training Az	Test Az
LDA Classifier based on temporal pairs	0.87±0.03	0.83±0.03
SVM Classifier based on temporal pairs	0.87±0.03	0.82±0.03
NN Classifier based on temporal pairs	0.87±0.03	0.84±0.03
LDA Classifier based on current images alone	0.82±0.03	0.76±0.04
MQSA radiologist	-	0.72±0.04

For the current classifier that was classifying 11 clusters without detection on the prior, an average of 6 features from the current images were selected (Table 5). The selected features included 1 current RLS feature, 3 current morphological features and 2 GLDS current features. All the features were consistently selected for all the training partitions. The test A_z by the current classifier was 0.72±0.14 for classifying the clusters as malignant or benign.

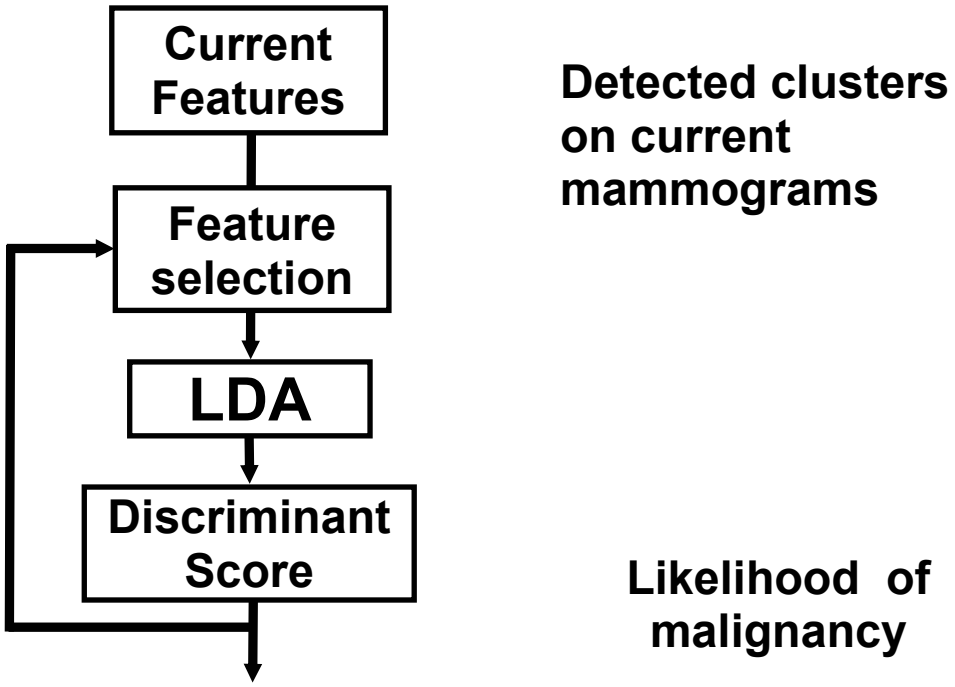

Figure 6. Current classifier based on current information alone. It is used in the case if no prior cluster is detected.

Table 5. Selected features for the combined classifier for classification on malignant and benign of automatically detected clusters in current and prior mammograms.

Classifier	Feature type	Ratio features	Difference features	Current features
Current classifier	Mo			3
	RLSF			1
	GLDS			2

In the future year we will continue to develop, improve and compare classification approaches in order to classify the clusters on malignant and benign based on automatically detected clusters and individual microcalcifications.

(C) Compare the classification accuracy of the classification scheme using temporal change information with that of a classifier using single-exam information alone. (Task 7).

We have performed a comparison between the temporal classifier and classifier based on current images (Table 6). The test A_z for the classifier based on 164 current images was 0.76±0.04. We see an improvement for the temporal classifier (A_z=0.83±0.03) when compared to the current classifier. We will continue to perform comparisons for the different design of both the temporal pair classifiers and the current classifiers.

Table 6. Results for the combined classifier for classification on malignant and benign of automatically detected clusters in current and prior mammograms.

Classifier Type	Temporal classifier	Current classifier
No. of pairs (or only current images)	164	11
Number of selected features	7	6
A_z	0.83±0.03	0.76±0.04

(D) Conduct observer performance study to compare radiologists' classification of malignant and benign microcalcifications with and without the aid of the temporal change classifier. (Task 8).

We started performing a pilot study as a first step to design an observer performance experiment with ROC methodology to evaluate the effects of computer classification on radiologists' estimates of the likelihood of malignancy of temporal pairs of microcalcification clusters [11, 12]. A graphical user interface was developed on a PC to display side-by-side the

temporal pairs of masses in a predesigned random order for each observer. The likelihood of malignancy and the BI-RADS assessment of the radiologist on each pair is automatically recorded when they select it on a slider.

175 pairs of serial mammograms from 65 patients containing clustered microcalcifications (51 malignant and 124 benign) were chosen from patient files. The true cluster locations were identified by an MQSA radiologist on all current mammograms and 164 of the priors. Regions of interest containing the corresponding clusters were extracted from the current and prior mammograms of each pair and analyzed by the CAD system. All cases eventually underwent biopsy so that interval change was observed for most of the clusters even if they were found to be benign after biopsy. This was therefore a difficult data set for interval change analysis.

Four MQSA radiologists, different from the one who identified the clusters, assessed the temporal pairs that were displayed side-by-side on a workstation. The cases were read in a counter-balanced design. The readers provided estimates of the LM and BI-RADS assessment without and then with CAD. The classification accuracy was quantified by the area under ROC curve, A_z.

The CAD system alone achieved a test Az value of 0.83 for this data set. For the four radiologists, the average Az in estimating the LM was 0.70 (range: 0.64 – 0.75) without CAD and improved to 0.77 (range: 0.68 – 0.83) with CAD. The improvement was statistically significant ($p=0.04$). The ROC study is ongoing to include additional radiologists as observers.

This pilot study indicates that CAD using interval change analysis can significantly improve radiologists' accuracy in classification of clustered microcalcifications on serial mammograms and it may be useful for assisting radiologists in classification of masses and thereby reducing unnecessary biopsies.

This pilot study will be the basis for our design of a full-scale ROC study, were the additional radiologists will provide estimates of the likelihood of malignancy and BI-RADS assessment without and with CAD. We are in process of recruiting additional breast radiologists at our department to participate as observers. The sample size is acceptable but we are continuing to enlarge the data set until the ROC study design is finalized. We also are working on the further improvement of the accuracy of the interval change classifier. We expect that this ROC study can be completed within the no cost time extension year approved for this grant. This type of observer study is new and unique and the outcome will be important, providing a new understanding of the potentials of computer aid to the radiologists in characterization of the temporal changes of mammographic masses.

(6) Key research accomplishments in current year as a result of this grant

- Increase of the temporal microcalcification database (collection of new temporal cases and extraction of regions of interest) (Task 1).
- Development of LDA, SVM and NN classifiers for classification on malignant and benign interval clusters based on automatically detected clusters and individual microcalcifications (Task 6).
- Comparison of temporal classifier and classifier based on current microcalcification clusters (Task 7).
- Performing a pilot observer study with the radiologist evaluating temporal pairs of microclacification clusters without CAD (Task 8).

(7) Reportable Outcomes

Publications in current year as a result of this grant

[1] Hadjiiski LM, HP Chan, B Sahiner, MA Helvie, MA Roubidoux, C Zhou, "Automated regional registration and classification of corresponding microcalcification clusters on serial mammograms," *91th Scientific Assembly and Annual Meeting of the Radiological Society of North America*, Chicago, IL, Nov. 27-Dec 2, 2005. pp. 270.

[2] Hadjiiski L, D Drouillard, HP Chan, B Sahiner, MA Helvie, MA Roubidoux, C Zhou, "Characterization of corresponding microcalcification clusters on temporal pairs of mammograms for interval change analysis: comparison of classifiers" *Proc. SPIE Medical Imaging,* 2006, 6144: 5Q1-5Q6.

[3] Hadjiiski L, HP Chan, B Sahiner, MA Roubidoux, MA Helvie, CParamagul, AV Nees, J Bailey, SK Patterson, "ROC Study: Effects of Computer-Aided Diagnosis on Radiologists' Characterization of Malignant and Benign Breast Clustered Microcalcifications in Temporal Pairs of Mammograms," To be presented at the *92nd Scientific Assembly and Annual Meeting of the Radiological Society of North America,* Chicago, IL. Nov. 26 – Dec. 1, 2006.

[4] Hadjiiski L, HP Chan, B Sahiner, AV Nees, J Bailey, SK Patterson, MA Roubidoux, MA Helvie, C Paramagul, "Computer-Aided Characterization of Malignant and Benign Breast Clustered Microcalcifications on Serial Mammograms: Early Experience of its Effects on Radiologists' Performance from an ROC Study," Education exhibit at the *92nd Scientific Assembly and Annual Meeting of the Radiological Society of North America,* Chicago, IL. Nov. 26 – Dec. 1, 2006.

Copies of publications are enclosed with this report.

(8) Conclusion

As a result of the support by the USAMRMC grant, in the fourth year of this project, we have (1) collected additional cases with temporal microcalcification clusters; (2) Extracted texture and morphological features from matching corresponding microcalcification clusters on current and prior mammograms from automatically detected microcalcification clusters; (3) Develop feature measures and temporal classifier for characterization of temporal changes in automatically detected microcalcification clusters; (4) Compare the performance of LDA, SVM and NN temporal classifiers (5) Compare the temporal classifier with classifier based on current images only; (6) Perform a pilot observer performance study.

The results obtained so far are encouraging. The new classification scheme, using interval change information, to classify mammographic microcalcification clusters as malignant and benign showed promising results. Morphological ratio features, difference RLS features and current GLDS features were useful for the classification.

We have made a major progress in the direction of the ultimate goal of the project - to have an automatic CAD system for characterization of temporal cases on malignant and benign. For that purpose a feature extraction and classification was carried out with the clusters and individual microcalcifications obtained automatically from the registration stage. Based on these features we have designed a classification scheme to classify the automatically registered and detected microcalcification clusters. This temporal classifier performed better than the classifier based on current images only.

We performed a pilot observer study for evaluation of the radiologist performance without and with CAD. Four radiologists participated in the study. The radiologists improved their performance when using CAD. The improvement was statistically significant.

The design of an automatic CAD can also present challenges. The automatically detected cluster locations on the temporal pairs may deviate from the optimal locations as selected by expert radiologists. This will introduce “noise” to the extracted features and make the classification task more difficult. In the feature year we will continue to improve all the stages of the automatic detection and classification system with the aim to improve the classification results for microcalcification clusters obtained automatically from the registration stage.

(9) References

1. L. Hadjiiski, H.P. Chan, B. Sahiner, C Zhou, M.A. Helvie, M.A. Roubidoux, "Computerized Regional Registration of Corresponding Masses and Microcalcification Clusters on Temporal Pairs of Mammograms for Interval Change Analysis", *89th Scientific Assembly and Annual Meeting of the Radiological Society of North America (RSNA)*, Chicago, Illinois, 2003.
2. Haralick RM, Shanmugam K and Dinstein I, "Texture features for image classification," IEEE Trans Sys Man and Cybern SMC-3, 610-621 (1973).
3. Sahiner B, Chan HP, Petrick N, Wei D, Helvie MA, Adler DD and Goodsitt MM, "Classification of mass and normal breast tissue: A convolution neural network classifier with spatial domain and texture images," IEEE Trans Med Img 15, 598-610 (1996).
4. Weszka JS, Dyer CR and Rosenfeld A, "A comparative study of texture measures for terrain classification," IEEE Trans Sys Man and Cybern 6, 269-285 (1976).
5. Filev P, LM Hadjiiski, B Sahiner, H-P Chan, MA Helvie, "Comparison of similarity measures for the task of template matching of masses on serial mammograms," *Med Phys* 2005, 32 (2), 515-529.
6. Hadjiiski LM, HP Chan, B Sahiner, MA Helvie, MA Roubidoux, C Zhou, "Interval change analysis based on computerized regional registration of corresponding microcalcification clusters on temporal pairs of mammograms," *90th Scientific Assembly and Annual Meeting of the Radiological Society of North America*, Chicago, IL, Nov. 28-Dec 3, 2004, pp. 491.
7. L. Hadjiiski, H.P. Chan, M. Gurcan, B. Sahiner, N. Petrick, M.A. Helvie, M. Roubidoux "Computer-Aided Characterization of Malignant and Benign Microcalcification Clusters Based on the Analysis of Temporal Change of Mammographic Features", Presented at the *SPIE International Symposium on Medical Imaging*, San Diego, California, February 23-28, 2002. *Proc. SPIE Medical Imaging, 2002, 4684, pp.749-753*.
8. V. N. Vapnik, *Statistical Learning Theory*, Wiley, New York, 1998.
9. Hadjiiski LM, HP Chan, B Sahiner, MA Helvie, MA Roubidoux, C Zhou, "Automated regional registration and classification of corresponding microcalcification clusters on serial mammograms," *91th Scientific Assembly and Annual Meeting of the Radiological Society of North America*, Chicago, IL, Nov. 27-Dec 2, 2005. pp. 270.
10. Hadjiiski L, D Drouillard, HP Chan, B Sahiner, MA Helvie, MA Roubidoux, C Zhou, "Characterization of corresponding microcalcification clusters on temporal pairs of mammograms for interval change analysis: comparison of classifiers" *Proc. SPIE Medical Imaging,* 2006, 6144: 5Q1-5Q6.

11. Hadjiiski L, HP Chan, B Sahiner, MA Roubidoux, MA Helvie, CParamagul, AV Nees, J Bailey, SK Patterson, "ROC Study: Effects of Computer-Aided Diagnosis on Radiologists' Characterization of Malignant and Benign Breast Clustered Microcalcifications in Temporal Pairs of Mammograms," To be presented at the *92nd Scientific Assembly and Annual Meeting of the Radiological Society of North America,* Chicago, IL. Nov. 26 – Dec. 1, 2006.
12. Hadjiiski L, HP Chan, B Sahiner, AV Nees, J Bailey, SK Patterson, MA Roubidoux, MA Helvie, C Paramagul, "Computer-Aided Characterization of Malignant and Benign Breast Clustered Microcalcifications on Serial Mammograms: Early Experience of its Effects on Radiologists' Performance from an ROC Study," Education exhibit at the *92nd Scientific Assembly and Annual Meeting of the Radiological Society of North America,* Chicago, IL. Nov. 26 – Dec. 1, 2006.

(10) Appendix

Copies of publications are enclosed with this report.

RSNA2005
Connecting for Lifelong Learning

Radiological Society of North America

Scientific Assembly and Annual Meeting Program

Radiological Society of North America
91st Scientific Assembly and Annual Meeting
November 27 – December 2, 2005
McCormick Place, Chicago

RSNA® Radiological Society of North America Founded in 1915

American Association of Physicists in Medicine

Invited Papers from ASNR, ASTRO, SIR and SNM
Search the RSNA Meeting Program at rsna2005.rsna.org

tion. Also, for each observer, sensitivity and specificity were determined from all 3 types of interpretations. Paired t-tests were performed to determine the significance of differences in the average sensitivities and specificities without and with CAD.
RESULTS: For both BIRADS and CR, it was not possible for every observer to choose a cutoff that resulted in a PMD that was in agreement with the actual PMD. However, use of user-dependent cutoffs resulted in better agreement than use of constant cutoff. Without the use of CAD, the averages over all observers of the sensitivity/specificity pair as determined from PMD, BIRADS, CR (2% cutoff), CR (a user dependent cutoff) were 0.88/0.66, 0.86/0.70, 1.00/0.03, and 0.93/0.46, respectively. With the use of CAD, the average sensitivity/specificity pairs were 0.93/0.67, 0.92/0.77, 1.00/0.04 and 0.95/0.56, respectively.
CONCLUSIONS: For many users, different forms of radiologists' interpretations result in different patient management decisions. Different forms of radiologists' interpretations may also result in different conclusions about the ability of CAD to improve the radiologist performance. Finally, when determining patient management from confidence ratings, user-dependent cutoffs are needed. (K.J.H.,M.L.G.,C.E.M.: Shareholders in R2 Technology, Inc. (Los Altos, CA).)

SSC17-05 • 11:10 AM

Computer-aided Diagnosis Scheme for Identifying Histological Classification of Clustered Microcalcifications Using Follow-up Magnification Mammograms

R. Nakayama, PhD★, Tsu, Japan • R. Watanabe, MD★ • K. Namba, MD★ • K. Takeda, MD★ • S. Katsuragawa, PhD • K. Doi, PhD • et al (nakayama@clin.medic.mie-u.ac.jp)

PURPOSE: The histological classification of clustered microcalcification on mammograms can be difficult, and thus often requires follow-up. The purpose of this study was to develop a computer-aided diagnosis scheme for identifying histological classification of clustered microcalcifications on follow-up magnification mammograms in order to assist radiologists' interpretation as a "second opinion."
METHOD AND MATERIALS: Our database consisted of 186 magnification mammograms which included current and previous images obtained from 93 patients after three months follow-up. It included 11 invasive carcinomas, 19 noninvasive carcinomas of comedo type, 25 noninvasive carcinomas of noncomedo type, 23 mastopathies and 15 fibroadenomas. The histological classifications of all clustered microcalcifications were proved by pathologic diagnosis after three months follow-up. The clustered microcalcifications were first segmented by use of a filter bank and a thresholding technique. Five objective features on clustered microcalcifications were determined from each of current and previous images by taking into account subjective features that experienced radiologists commonly use to identify possible histological classifications. Bayes discriminant function was employed for distinction between five histological classifications. For the input of Bayes discriminant function, we used five objective features obtained from current images, and also ten objective features obtained from both current and previous images.
RESULTS: With current images, the classification accuracy ranged from 76.0% to 90.9%, but was improved substantially by use of previous images. With ten objective features, the classification accuracies of our computerized scheme were 100%(11/11) for invasive carcinoma, 100%(19/19) for noninvasive carcinoma of comedo type, 96.0%(24/25) for noninvasive carcinoma of noncomedo type, 95.7% (22/23) for mastopathy, and 100% (15/15) for fibroadenoma.
CONCLUSIONS: By use of both current and previous images, this computerized scheme achieved high classification accuracies, and would be useful to assist radiologists in their assessment of clustered microcalcifications. (S.K, K.D. are shareholders in R2 Technology.)

SSC17-06 • 11:20 AM

Automated Regional Registration and Classification of Corresponding Microcalcification Clusters on Serial Mammograms

L.M. Hadjiiski, PhD★, Ann Arbor, MI • H. Chan, PhD★ • B. Sahiner, PhD★ • M.A. Helvie, MD★ • M.A. Roubidoux, MD★ • C. Zhou, PhD★ (lhadjiisk@umich.edu)

PURPOSE: To develop an automated system for detecting corresponding microcalcification clusters on serial mammograms, and classifying the cluster as malignant and benign using interval change information.
METHOD AND MATERIALS: Our system consists of two stages. In the first stage, based on the location of a detected cluster on the current mammogram, a regional registration procedure identifies the local area on the prior that may contain the corresponding cluster. A search program is used to detect cluster candidates within the local area. The cluster on the current image is then paired with the candidates to form true (TP-TP) or false (TP-FP) pairs. A correspondence classifier using automatically extracted features is designed to reduce the false pairs. In the second stage, a temporal classifier based on current and prior information is used if a cluster is detected in the prior, and a current classifier based on current information alone is used if no prior cluster is detected. In this study, 175 serial pairs containing biopsy-proven calcification clusters were used. An MQSA radiologist identified the corresponding clusters on the mammograms. On the priors, the radiologist rated 12 of the 175 clusters as not visible and the subtlety of 18 clusters as 9 or 10 on a scale of 10. Leave-one-case-out resampling was used for feature selection and classification.
RESULTS: The search program detected 90.2% (158/175) of the clusters on the priors with an average of 0.43 FPs/image. The correspondence classifier identified 85% (149/175) of the TP-TP pairs with 15 false matches within the 164 image pairs that had detected clusters. In the classification stage the temporal classifier achieved a test Az of 0.83 for the 164 pairs for classifying the clusters as malignant or benign. For the 11 clusters without detection on the prior, the test Az by the current classifier was 0.72. In comparison, the MQSA radiologist achieved an Az of 0.72 for both the 175 and the 164 temporal pairs.
CONCLUSIONS: Our interval change analysis system can detect the corresponding cluster on the prior mammogram with high sensitivity, and classify them with an accuracy comparable to that an experienced radiologist.

SSC17-07 • 11:30 AM

Workflow Analysis for a Clinical Breast Ultrasound CAD Workstation Prototype

N. Gruszauskas★, Chicago, IL • K. Drukker, PhD★ • C.A. Sennett, MD★ • L. Lan, MS★ • M.L. Giger, PhD • I. Bonta, MD★ (ngrusz1@uic.edu)

PURPOSE: To evaluate the impact of a prototype breast ultrasound CAD workstation on the workflow in a clinical environment, including its integration with standard RIS equipment and its use by radiologists.
METHOD AND MATERIALS: We designed a breast ultrasound CAD prototype workstation to be utilized in a clinical setting based on previous experience with retrospective image analysis using CAD methodology. In its current implementation, the workstation is configured as an archive destination on our clinical ultrasound scanner, an HDI 5000 by Philips Medical Systems. We designed an automatic image transfer method using DICOM protocols to transfer images directly from the scanner to the workstation. There, the images are automatically preprocessed and all patient-identifiable information is removed using in-house developed software. The radiologist then performs lesion identification on each image using our customized graphical user interface. For the purposes of this workflow study, CAD analysis results were not displayed in order to prevent influence of patient care. Using our prototype workstation, images from 40 patients – who consented to participate in this study – were obtained from exams performed by a single radiologist. The radiologist was surveyed to determine the workstation's overall impact on clinical workflow.
RESULTS: Our workstation design was able to completely interoperate with the clinical DICOM-compliant equipment. The amount of time necessary to transfer an entire exam to the workstation varied substantially. However, the radiologist reported that the workstation added only one to two minutes to the amount of time necessary to perform a standard ultrasound workup.
CONCLUSIONS: The use of our prototype workstation has no significant impact on the time required to perform a standard diagnostic breast ultrasound exam and on the clinical workflow in general. Our design allows for easy integration into a standard clinical environment due to its utilization of industry standard DICOM protocols. (M.L.G.: M. L. Giger is a shareholder in R2 Technology, Inc (Sunnyvale, CA). It is the policy of the University of Chicago that investigators disclose publicly actual or potential significant financial interests that may appear to be affected by the research activities.)

SSC17-08 • 11:40 AM

Volumetric Breast Density Estimation from Full Field Digital Mammograms

N. Karssemeijer, PhD★, Nijmegen, Netherlands • S. van Engeland★ • C. Boetes, MD★ • H. Huisman, PhD★ • P. Snoeren, PhD★

PURPOSE: Breast density is an important risk factor for breast cancer. The purpose of this paper is to present a method for accurate estimation of volume of dense breast tissue from full field digital mammograms (FFDM) and to validate results by correlation with volumes of dense breast tissue determined from breast MRI.
METHOD AND MATERIALS: Unprocessed FFDM images allow conversion of pixel values to the log exposure domain. By using a physical model of image acquisition this allows estimation of the proportion of dense breast tissue contributing to each pixel. By integration, the volume of dense tissue can be obtained. The physical model we use is based on the assumption that the breast is composed of two types of tissue, fatty and dense. Effective linear attenuation coefficients of these tissues were derived from empirical data as a function of three parameters: Tube voltage (kVp), anode material, and filtration. These parameters are available from the DICOM image headers. A simple relation is derived that allows computa-

★ - Author stated no financial disclosure ▲ - Disclosure information unavailable at press time

Characterization of Corresponding Microcalcification Clusters on Temporal Pairs of Mammograms for Interval Change Analysis - Comparison of Classifiers

Lubomir Hadjiiski*, Douglas Drouillard, Heang-Ping Chan, Berkman Sahiner, Mark A. Helvie, Marilyn Roubidoux, Chuan Zhou

Department of Radiology, University of Michigan, Ann Arbor, MI 48109

ABSTRACT

We are developing an automated system for analysis of microcalcification clusters on serial mammograms. Our automated system consists of two stages: (1) automatic registration of corresponding clusters on temporal pairs of mammograms producing true (TP-TP) and false (TP-FP) pairs; and (2) characterization of temporal pairs of clusters as malignant and benign using a temporal classifier. In this study, we focussed on the design of the temporal classifier. Morphological and texture (RLS and GLDS) features are automatically extracted from the detected current and prior cluster locations. Additionally, difference morphological and RLS features are obtained. The automatically detected cluster locations on the temporal pairs may deviate from the optimal locations as selected by expert radiologists. This will introduce "noise" to the extracted features and make the classification task more difficult. Linear discriminant analysis (LDA) and support vector machine (SVM) classifiers were trained to classify the true and false pairs. Leave-one-case-out resampling method was used for feature selection and classifier design. In this study, 175 serial mammogram pairs containing biopsy-proven microcalcification clusters were used. At the first stage of the system, 85% (149/175) of the TP-TP pairs were identified with 15 false matches within the 164 image pairs that had computer-detected clusters on the priors. At the second stage, an average of 7 features were selected (4 difference morphological, 1 difference RLS and 2 current GLDS). The LDA and SVM temporal classifiers achieved test A_z of 0.83 and 0.82, respectively, for the classification of the 164 cluster temporal pairs as malignant or benign. In comparison, an MQSA radiologist achieved an A_z of 0.72. Both the LDA and SVM classifiers were able to classify the automatically detected temporal pairs of microcalcification clusters with accuracy comparable to that of an experienced radiologist.

Keywords: Computer-Aided Diagnosis, Interval Changes, Microcalcification Classification, Feature analysis, Mammography, Malignancy.

1. INTRODUCTION

Mammography is currently the most effective method for early breast cancer detection[1,2]. Radiologists routinely compare mammograms from a current examination with those obtained in previous years, if available, for identifying interval changes, detecting potential abnormalities, and evaluating breast lesions. It is widely accepted that analysis of interval changes in mammographic features is very useful for both detection and classification of abnormalities[3,4]. A variety of computer-aided diagnosis (CAD) techniques have been developed to detect mammographic abnormalities and to distinguish between malignant and benign lesions. We are studying the use of CAD techniques to assist radiologists in interval change analysis.

Commonly used classification methods for CAD use information from a single examination. These methods have been shown to perform well in lesion classification problems[5-14]. However, when multiple-year mammograms of a

* L. H. (correspondence): e-mail: lhadjiski@umich.edu

Medical Imaging 2006: Image Processing, edited by Joseph M. Reinhardt, Josien P. W. Pluim,
Proc. of SPIE Vol. 6144, 61445Q, (2006) · 0277-786X/06/$15 · doi: 10.1117/12.654920

lesion are available, it is not trivial to design computer vision methods to use the temporal information for computer-aided classification and to improve the differentiation between benign and malignant masses.

The goal of our research is to develop a technique for computerized analysis of temporal differences between a microcalcification cluster on the most recent mammogram and a prior mammogram of the same view[15]. The computer system can be used to assist radiologists in evaluating interval changes and thus distinguishing between malignant and benign microcalcification clusters for CAD. It will also be useful for improving the identification of new or growing clusters in a detection system or for improved classification of malignant and benign clusters in a diagnostic system. In our previous studies we have demonstrated that interval change analysis can improve differentiation of malignant and benign masses[16,17].

The purpose of this study is to compare the performance of different classifiers for the task of automated characterization of microcalcification clusters using interval change information on serial mammograms. With an automated system, noise is introduced by the imperfect registration and imperfect detection of the clusters on both current and prior mammograms. This analysis will be useful for understanding the advantages and limitations of the classifiers when performing in noisy conditions.

2. MATERIALS AND METHODS

Our automated system consists of two stages: (1) registration of corresponding clusters on temporal pairs of mammograms, and (2) characterization of the temporal pairs of clusters as malignant and benign (Fig. 1). A description of the detection method and the database used is given below.

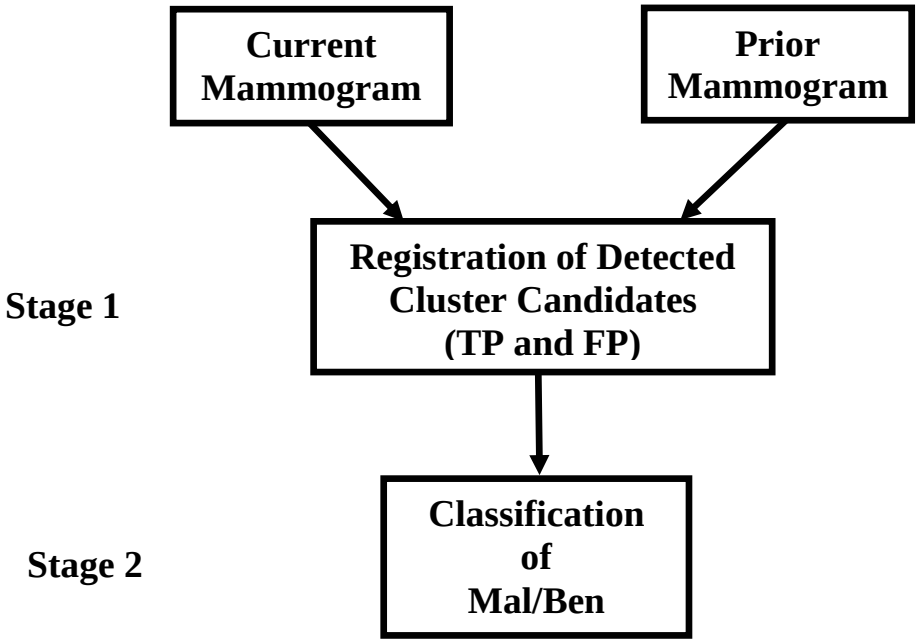

Figure 1. Block-diagram of the registration – classification CAD for temporal microcalcification clusters.

2.1 Registration of corresponding microcalcification clusters in serial mammograms

In the first stage, based on the location of a detected cluster on the current mammogram, a regional registration procedure identifies the local area on the prior that may contain the corresponding cluster. An automated detection program is used to detect cluster candidates within the local area (Fig. 2) that may include true positives (TP) and false positives (FPs). The cluster on the current image is then paired with the candidates to form true (TP-TP) or false (TP-FP) pairs (Fig. 2). A correspondence classifier is used to reduce the false pairs (TP-FP). This technique was presented in detail previously[18].

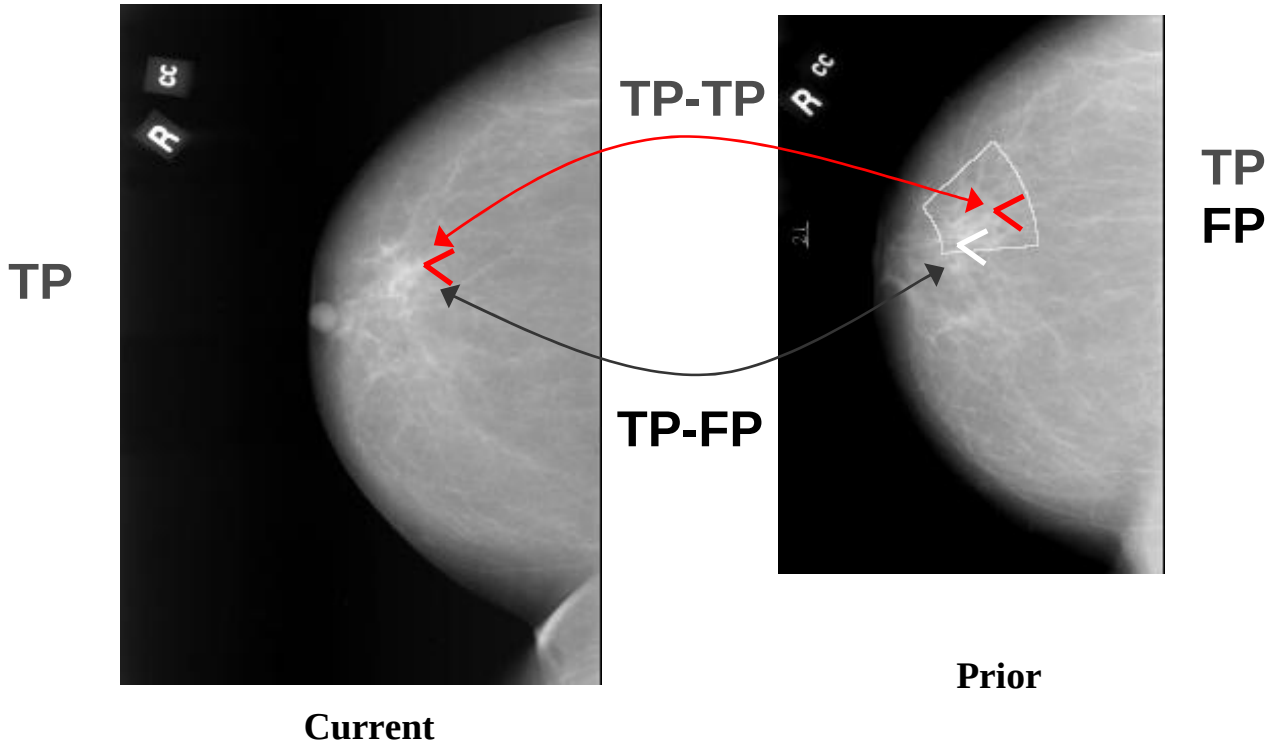

Figure 2. Automated registration resulting in true (TP-TP) and false (TP-FP) pairs by pairing the cluster on the current image with the detected TP and FP candidates on prior image.

2.2 Classification of automatically detected and registered temporal pairs of microcalcification clusters

In the second stage, a temporal classifier based on current and prior information is used if a cluster is detected on the prior, and a current classifier based on current information alone is used if no prior cluster is detected. In this study, we focussed on the design of the temporal classifier (Fig. 3). Morphological and texture features including run length statistics (RLS) and gray level dependance statistics (GLDS) are automatically extracted from the detected current and prior cluster locations. Sixty four GLDS features, 20 morphological features and 20 RLS features and the cluster size were extracted. Additionally, 20 difference morphological and 20 difference RLS features are obtained by subtracting a prior feature from the corresponding current feature. The automatically detected cluster locations on the temporal pairs may deviate from the optimal locations as selected by expert radiologists. This will introduce "noise" to the extracted features and make the classification task more difficult. Linear discriminant analysis (LDA) and support vector machine[19] (SVM) classifiers were trained to classify the true and false pairs.

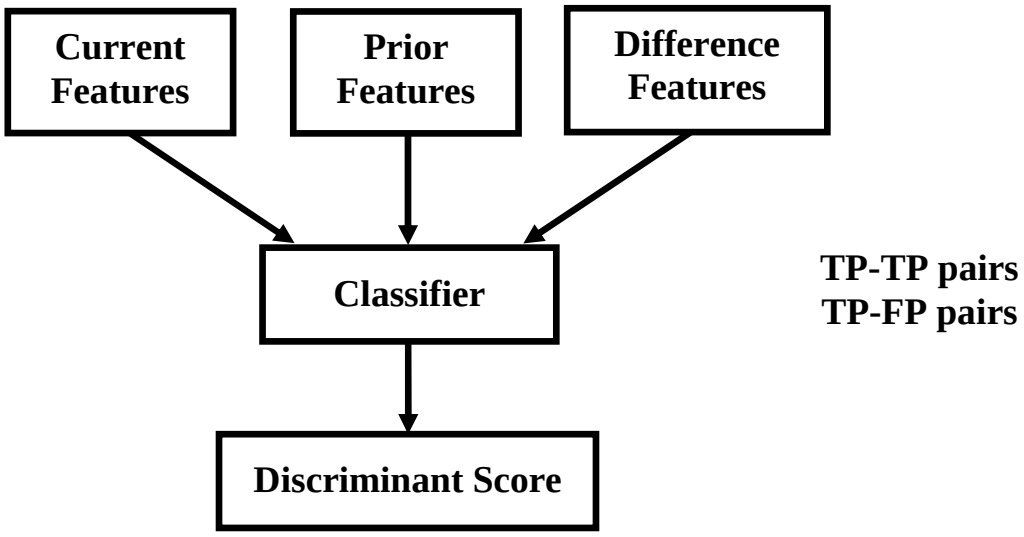

Figure 3. Block-diagram of the temporal classifier for classification of malignant and benign microcalcification clusters. Both TP-TP and TP-FP pairs can be input to the system.

2.3 Data set

In this study, 175 serial mammogram pairs containing biopsy-proven microcalcification clusters were used, of which 51 were biopsy-proven to be malignant and 124 benign. An experienced MQSA radiologist identified the gold standard cluster locations on corresponding mammogram pairs. On the priors, the radiologist rated 12 of the 175 clusters as not visible and the subtlety of 18 clusters as 9 or 10 on a scale of 10 (1=most obvious, 10=subtlest).

The mammograms were digitized with a LUMISCAN 85 laser scanner at a pixel resolution of $50\mu m \times 50\mu m$ and 4096 gray levels. The image matrix size was reduced by averaging every 2 x 2 adjacent pixels and down-sampled by a factor of 2 to obtain images with a pixel size of $100\mu m \times 100\mu m$ for analysis by the computer.

2.4 ROC analysis

Leave-one-case-out resampling method was used for feature selection and classifier design. Stepwise feature selection was used to obtain the best feature set. To evaluate the classifier performance, the training and test discriminant scores were analyzed using receiver operating characteristic (ROC) methodology[20]. The discriminant scores of the malignant and benign microcalcification clusters were used as decision variables in the LABROC1 program[21], which fits a binormal ROC curve based on maximum likelihood estimation. The classification accuracy was evaluated as the area under the ROC curve, A_z.

3. RESULTS

At the first stage of the system, 85% (149/175) of the TP-TP pairs were identified with 15 false matches within the 164 image pairs that had computer detected clusters on the priors. Therefore 164 temporal image pairs (149 TP-TP and 15 TP-FP), of which 49 were malignant, were used for classification as malignant and benign. At the second stage, an average of 7 features were selected. The selected features included 4 difference morphological features, 1 difference RLS feature and 2 GLDS features from the current image. The SVM and LDA temporal classifiers were used for the classification of the 164 cluster temporal pairs. SVMs with dot kernel, radial kernel and tanh kernel (neural kernel) were

studied. The training and testing results are presented in Table 1. The SVM classifier achieved test A_z of 0.82±0.03. The LDA classifier achieved test A_z of 0.83±0.03. In comparison, the MQSA radiologist achieved an A_z of 0.72±0.04 for the 164 temporal pairs. The difference in A_z between any one of the classifiers and the radiologist did not achieve statistical significance.

Table 1. SVM classification results based on 7 selected features.

SVM Type	Training A_z	Test A_z	Number SV
Dot Kernel	0.87±0.03	0.82±0.03	78
Radial Kernel	0.86±0.03	0.81±0.03	89
Tanh Kernel	0.86±0.03	0.82±0.03	83

4. CONCLUSION

We studied the use of the LDA and SVM classifiers for the classification of automatically detected temporal pairs of microcalcification clusters as malignant or benign. Both the LDA and SVM classifiers were able to classify the automatically detected temporal pairs of microcalcification clusters with an accuracy comparable to that of an experienced radiologist. The SVMs with dot kernel, radial kernel, and tanh kernel had comparable performance with the radial kernel being slightly worse than the other two. The LDA and SVM classifiers also had comparable performance.

ACKNOWLEDGMENTS

This work was supported by USAMRMC grants DAMD17-02-1-0489. The authors are grateful to Charles E. Metz, Ph.D., for the LABROC program.

REFERENCES

1. S. A. Feig, C. J. D'Orsi, R. E. Hendrick, V. P. Jackson, D. B. Kopans, B. Monsees, E. A. Sickles, C. B. Stelling, M. Zinninger, and P. Wilcox-Buchalla, "American College of Radiology guidelines for breast cancer screening," *AJR Am J Roentgenol.* **171**, 29-33, 1998.

2. B. Cady and J. S. Michaelson, "The life-sparing potential of mammographic screening," *CANCER* **91**, 1699-1703, 2001.

3. L. W. Bassett, B. Shayestehfar, and I. Hirbawi, "Obtaining previous mammograms for comparison: usefullness and costs," *American Journal of Roentgenology* **163**, 1083-1086, 1994.

4. E. S. Burnside, E. A. Sickles, R. E. Sohlich, and K. E. Dee, "Differential Value of Comparison with Previous Examinations in Diagnostic Versus Screening Mammography," *Amer. J. Roentgenology* **179**, 1173-1177, 2002.

5. H. P. Chan, D. Wei, K. L. Lam, B. Sahiner, M. A. Helvie, D. D. Adler, and M. M. Goodsitt, "Classification of malignant and benign microcalcifications by texture analysis," *Medical Physics* **22**, 938, 1995.

6. Y. Jiang, R. M. Nishikawa, D. E. Wolverton, C. E. Metz, M. L. Giger, R. A. Schmidt, C. J. Vyborny, and K. Doi, "Malignant and benign clustered microcalcifications: automated feature analysis and classification," *Radiology* **198**, 671-678, 1996.

7. H. P. Chan, B. Sahiner, N. Petrick, M. A. Helvie, K. L. Leung, D. D. Adler, and M. M. Goodsitt, "Computerized classification of malignant and benign microcalcifications on mammograms: texture analysis using an artificial neural network," *Physics in Medicine and Biology* **42**, 549-567, 1997.

8. H. P. Chan, B. Sahiner, K. L. Lam, N. Petrick, M. A. Helvie, M. M. Goodsitt, and D. D. Adler, "Computerized analysis of mammographic microcalcifications in morphological and texture feature space," *Medical Physics* **25**, 2007-2019, 1998.

9. Y. Jiang, R. M. Nishikawa, and J. Papaioannou, "Dependence of computer classification of clustered microcalcifications on the correct detection of microcalcifications " *Medical Physics* **28**, 1949-1957, 2001.

10. M. F. Salfity, R. M. Nishikawa, Y. Jiang, and J. Papaioannou, "The use of a priori information in the detection of mammographic microcalcifications to improve their classification," *Medical Physics* **30**, 823-831, 2003.

11. S. Paquerault, L. M. Yarusso, J. Papaioannou, Y. Jiang, and R. M. Nishikawa, "Radial gradient-based segmentation of mammographic microcalcifications: Observer evaluation and effect on CAD performance," *Medical Physics* **31**, 2648-2657, 2004.

12. M. Kallergi, "Computer-aided diagnosis of mammographic microcalcification clusters," *Medical Physics* **31**, 314-326, 2004.

13. I. Leichter, R. Lederman, S. S. Buchbinder, P. Bamberger, B. Novak, and S. Fields, "Computerized Evaluation of Mammographic Lesions: What Diagnostic Role Does the Shape of the Individual Microcalcifications Play Compared with the Geometry of the Cluster?," *AJR Am J Roentgenol.* **2004**, 705–712, 2004.

14. L. Wei, Y. Yang, R. M. Nishikawa, and Y. Jiang, "A Study on Several Machine-Learning Methods for Classification of Malignant and Benign Clustered Microcalcifications," *IEEE Trans Medical Imaging* **24**, 371-380, 2005.

15. L. M. Hadjiiski, H. P. Chan, B. Sahiner, N. Petrick, M. A. Helvie, M. A. Roubidoux, and M. N. Gurcan, "Computer-aided characterization of malignant and benign microcalcification clusters based on the analysis of temporal change of mammographic features," *Proc SPIE* **4684**, 749-753, 2002.

16. L. M. Hadjiiski, B. Sahiner, H. P. Chan, N. Petrick, M. A. Helvie, and M. N. Gurcan, "Analysis of Temporal Change of Mammographic Features: Computer-Aided Classification of Malignant and Benign Breast Masses," *Medical Physics* **28**, 2309-2317, 2001.

17. L. M. Hadjiiski, H. P. Chan, B. Sahiner, M. A. Helvie, M. Roubidoux, C. Blane, C. Paramagul, N. Petrick, J. Bailey, K. Klein, et al., "Improvement of Radiologists' Characterization of Malignant and Benign Breast Masses in Serial Mammograms by Computer-Aided Diagnosis: An ROC Study," *Radiology* **233**, 255-265, 2004.

18. L. M. Hadjiiski, H. P. Chan, B. Sahiner, M. A. Helvie, M. A. Roubidoux, and C. Zhou, "Interval Change Analysis Based on Computerized Regional Registration of Corresponding Microcalcification Clusters on Temporal Pairs of Mammograms," 491-491, RSNA Program Book 2004, 2004.

19. V. N. Vapnik, *Statistical Learning Theory*, Wiley, New York, 1998.

20. C. E. Metz, "ROC methodology in radiologic imaging," *Investigative Radiology* **21**, 720-733, 1986.

21. C. E. Metz, J. H. Shen, and B. A. Herman, "New methods for estimating a binormal ROC curve from continuously-distributed test results," *Annual Meeting of the American Statistical Association*, Anaheim, CA, 1990.

Abstract ID: 4442391

Submission Type: Scientific Papers

Submission Status: Accepted

Contact:
Lubomir Hadjiiski
University of Michigan

Phone: 734-647-7428
Fax: 734-615-5513
E-Mail: lhadjisk@umich.edu

Primary Category: Physics
Secondary Category: Image Processing, CAD, etc

04) ROC Study: Effects of Computer-aided Diagnosis on Radiologists' Characterization of Malignant and Benign Breast Clustered Microcalcifications in Temporal Pairs of Mammograms

L M Hadjiiski, PhD, Ann Arbor, MI; H Chan, PhD; B Sahiner, PhD; M A Roubidoux, MD; M A Helvie, MD; C P Paramagul, MD; et al. (lhadjisk@umich.edu)

PURPOSE

To evaluate the effects of computer-aided diagnosis (CAD) on radiologists' characterization of clustered microcalcifications on serial mammograms.

METHOD AND MATERIALS

We developed an automated CAD system to register and classify microcalcification clusters as malignant or benign based on interval change information on serial mammograms. In this study, we conducted observer performance experiments with receiver operating characteristic (ROC) methodology to evaluate the effects of CAD on radiologists' estimates of the likelihood of malignancy (LM) of the clusters. 175 pairs of serial mammograms from 65 patients containing clustered microcalcifications (51 malignant and 124 benign) were chosen from patient files with IRB approval. The true cluster locations were identified by an MQSA radiologist on all current mammograms and 164 of the priors. Regions of interest containing the corresponding clusters were extracted from the current and prior mammograms of each pair and analyzed by the CAD system. All cases eventually underwent biopsy so that interval change was observed for most of the clusters even if they were found to be benign after biopsy. This was therefore a difficult data set for interval change analysis. Four MQSA radiologists, different from the one who identified the clusters, assessed the temporal pairs that were displayed side-by-side on a workstation. The cases were read in a counter-balanced design. The readers provided estimates of the LM and BI-RADS assessment without and then with CAD. The classification accuracy was quantified by the area under ROC curve, Az.

RESULTS

The CAD system alone achieved a test Az value of 0.83 for this data set. For the four radiologists, the average Az in estimating the LM was 0.70 (range: 0.64–0.75) without CAD and improved to 0.77 (range: 0.68–0.83) with CAD. The improvement was statistically significant ($p=0.04$). The ROC study is ongoing to include additional radiologists as observers.

CONCLUSION

CAD using interval change analysis can significantly improve radiologists' accuracy in classification of clustered microcalcifications on serial mammograms.

CLINICAL RELEVANCE/APPLICATION

CAD using interval change analysis may increase the positive predictive value of mammography.

Disclosures:

No Disclosure:	Lubomir Hadjiiski
No Disclosure:	Heang-Ping Chan
No Disclosure:	Berkman Sahiner
No Disclosure:	Marilyn Roubidoux

No Disclosure:	Mark Helvie
No Disclosure:	Chintana Paramagul
No Disclosure:	Alexis Nees
No Disclosure:	Janet Bailey
No Disclosure:	Stephanie Patterson

Questions:

1. **Published email:** Do you wish to have an email address published in the RSNA program?

Yes

If yes, please provide one email address:

lhadjisk@umich.edu

2. **SUBMISSION OF MANUSCRIPT** An important part of scientific progress is the timely publication of new research; also important is the thorough peer review of all manuscripts prior to acceptance for publication. (To view how to submit your manuscript, click on (Publications) under the above Help menu). Please tell us to which of the following journals you plan to submit your work (select one):

No response

If you marked OTHER please indicate the name of the journal.

No response

3. **Digital Scientific Presentions**: With the permission of the authors, RSNA will post digital scientific presentations on the Web for a maximum of one year after the RSNA annual meeting. RSNA scientific presentations posted on the Web are not considered prior publication by the editors of *Radiology* and *Medical Physics*. Thus the manuscript of your presentation will be considered for publication if you choose to submit it.

No response

Abstract ID: 4441387

Submission Type: Education Exhibits

Submission Status: Accepted

Contact:
Lubomir Hadjiiski
University of Michigan

Phone: 734-647-7428
Fax: 734-615-5513
E-Mail: lhadjisk@umich.edu

Primary Category: Physics
Secondary Category: Diagnostic

Computer-Aided Characterization of Malignant and Benign Breast Clustered Microcalcifications on Serial Mammograms: Early Experience of Its Effects on Radiologists' Performance from an ROC Study

L M Hadjiiski, PhD, Ann Arbor, MI; H Chan, PhD; B Sahiner, PhD; A V Nees, MD; J E Bailey, MD; S K Patterson, MD; et al. (lhadjisk@umich.edu)

PURPOSE/AIM

(1) To understand the design of an observer performance study for characterization of microcalcification clusters on serial mammograms without and with CAD. (2) To understand the effects of CAD on radiologists' assessment of the likelihood of malignancy (LM) of clustered microcalcifications.

CONTENT ORGANIZATION

We will describe our CAD system that performs automated registration of corresponding clusters on serial mammograms and classifies malignant and benign clusters based on interval change information. We will describe the design of our multi-reader multi-case ROC study, and discuss the effects of CAD on radiologists' estimates of the likelihood of malignancy of microcalcifications using ROC analysis and on radiologists' biopsy recommendation decision by analysis of their BI-RADS assessment.

SUMMARY

CAD using interval change analysis may be useful for assisting radiologists in classification of microcalcification clusters and thereby reducing benign biopsies. This exhibit will lead to understanding of (1) Design of a CAD system using interval change analysis. (2) Design of an observer performance study for evaluation of CAD using receiver operating characteristic (ROC) methodology. (3) Interaction of observer with a CAD system (4) Effects of CAD on radiologists' decision.

Disclosures:

No Disclosure:	Lubomir Hadjiiski
No Disclosure:	Heang-Ping Chan
No Disclosure:	Berkman Sahiner
No Disclosure:	Alexis Nees
No Disclosure:	Janet Bailey
No Disclosure:	Stephanie Patterson
No Disclosure:	Marilyn Roubidoux
No Disclosure:	Mark Helvie
No Disclosure:	Chintana Paramagul

Questions:

1. **Published email:** Do you wish to have an email address published in the RSNA program?

Yes

If yes, please provide one email address:

lhadjisk@umich.edu

2. **Publication Policy:** Invitations to submit a manuscript to *RadioGraphics* are extended during RSNA week [see HELP-Education Exhibits]. The invitation will clearly indicate it is from

RadioGraphics. No other entity besides *RadioGraphics* is authorized to invite your exhibit for possible publication. RSNA requires that authors certify that no other organization has any rights to the material. RSNA reserves the right to accept or reject any abstract according to its usual and customary peer-review process. Please check the appropriate statement below:

[To avoid confusion, please check with the Lakeside Learning Center Help Desk in Hall D if an invitation from any other publisher, including a dot-com organization, is attached to your exhibit.]

No response

If yes, where:

No response

3. **No Recording Policy:** Attendees in the Lakeside Learning Center area of the RSNA Annual Meeting routinely photograph or capture in electronic format many of the presentations on display for their own individualized educational use. Photography (or other electronic recording) has been allowed with the understanding that the work presented is the property of the author; attendees may not reproduce, distribute, or publish any of the presentations captured at the RSNA annual meeting.

If 'No' response is selected below, a 'no recording' icon will be added to the presentation (on the sign above the hard-copy presentation and to the header of the online version). It is the meeting attendee's responsibility to abide by the 'no photography' designation. RSNA assumes no responsibility for meeting attendees who do not comply.

In recognition of intellectual property rights of authors, please indicate one of the following:

No response

4. **AV**: The meeting rooms and info**RAD** theater will be furnished with a PC computer (Pentium-4/1.3 GHz, 256MB RAM), video/data projector with 1024x786 resolution, mouse access for presenters at lectern, and a 15" flat LCD monitor next to lectern.

No response

If yes, please describe. NOTE: All special requirements are subject to review for approval.

No response